3-Heights™ PDF Analysis & Repair Service

Version 4.5

User Manual

Contact: pdfsupport@pdf-tools.com

Owner: **PDF Tools AG**

Kasernenstrasse 1
8184 Bachenbülach
Switzerland

http://www.pdf-tools.com

Copyright © 2000-2015

Table of Contents

1 Introduction 4

1.1 Description 4
1.2 Functions 4
Features 5
Formats 5
Compliance 5
1.3 Operating Systems 5

2 Installation 6

2.1 Overview 6
2.2 Installing the Service 6
2.3 Editing the Configuration File "pdfrepairsvr.ini" 6
Retrieve Information about Available Options 7
2.4 Creating and Starting the Service 8
2.5 Error Log Files 11
Log File per Service 11
Log File per Document 11
2.6 State Diagram of the Service 12
2.7 Difference between Repairable and Recoverable 13

3 License Management 14

3.1 Graphical License Manager Tool 14
List all installed license keys 14
Add and delete license keys 14
Display the properties of a license 14
Select between different license keys for a single product 15
3.2 Platform considerations 15
Windows 15

4 Reference Manual 15

4.1 Service Control Commands 15
-c Create Service 15
-d Delete Service 15
-s Start Service 16
-t Stop Service 16
-o Continue Service 16
-a Pause Service 16
-q Query Current Sstatus 16
-x Run as Executable 16
4.2 Configuration Options 16
Configuration File *Pdf2PdfSvr.ini* 17
-a Analyze Only 17
-b Set the Rebuild Options 17
-dp Do Not Recover Pages 18
-dx Do Not Recover Cross Reference Table 18
-p Read an Encrypted PDF File 18
-w Set the Watched Folder 18
-wd Set the Drop-In Folder 18
-wfs Process only files with certain extensions 19

-wfi Ignore files with certain extensions 19
Autodelete Successfully Processed Files 19
Job Number Prefix 19
Logpath 19
Polling Interval 20

5 Troubleshooting 20

5.1 The Repair Time Takes Too Long 20
5.2 The File Cannot be Repaired 20

© PDF Tools AG – Premium PDF Technology

1 Introduction

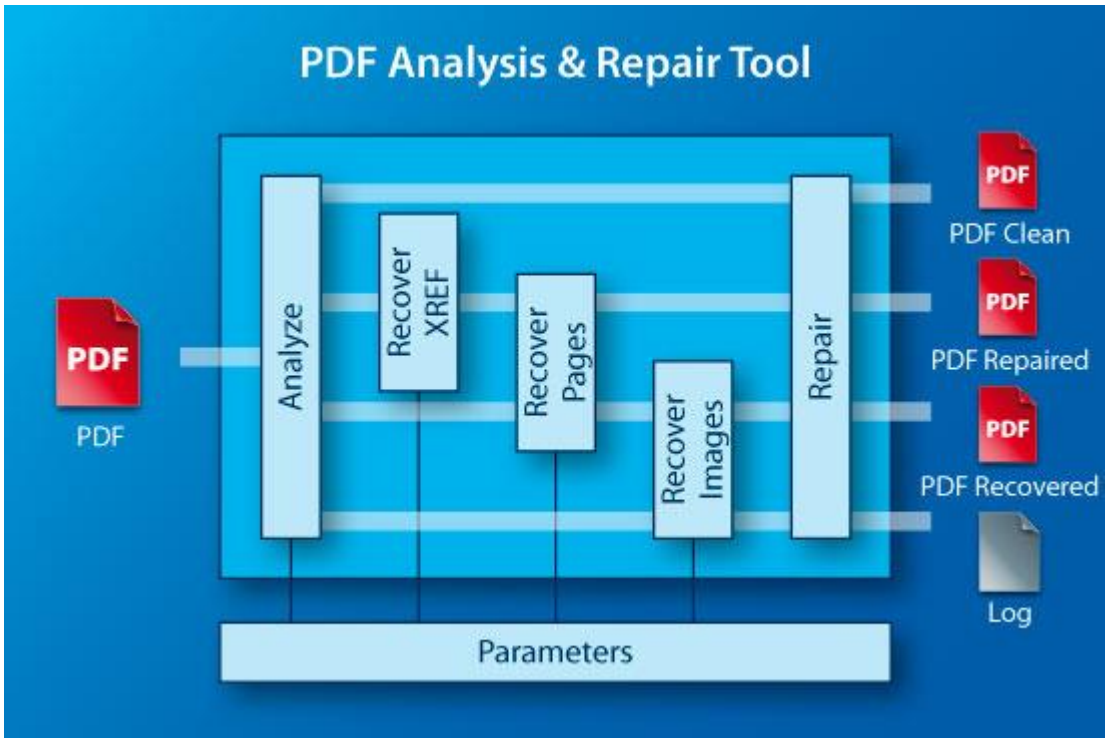
The 3-Heights™ PDF Analysis & Repair Service is a ready-to-use product that allows to install a Windows NT service process to automatically analyze and repair PDF documents from watched folders.

1.1 Description

Unfortunately, the number of corrupt PDF documents is incredibly huge. The cause is usually down to defective generating tools, converters and other influences such as attempts at manual editing, copying via FTP without correct settings, system crashes during PDF creation, network interruptions, defective copying on optical media, etc.

The result leads to an enormous loss of important information and to production downtimes caused by corrupt PDF documents.

The 3-Heights™ PDF Analysis & Repair analyzes PDF documents with regard to PDF specifications. Defective files are automatically repaired as far as possible and unreadable data is restored.



1.2 Functions

PDF Analysis & Repair is used to check and, where indicated, repair PDF documents. Users can determine customized profiles from a broad range of analysis and repair options. An exact and detailed description is issued for each reported error. The tool is also capable of reading and processing encrypted PDF files without any problems.

© PDF Tools AG – Premium PDF Technology

Features

Analyze and/or repair one or more PDF documents

Set analysis options, including:

- Objects
- Page tree
- Content stream

Set repair options, including:

- Restore data streams
- Restore fonts
- Restore XRef table
- Restore pages
- Restore images if pages cannot be restored

Display error description for every message, including:

- Type (errors, warnings, information)
- Error code
- Text-based description
- Page number
- Number of events

Write error messages to log file

Read encrypted PDF files

Encrypt restored file and set user authorizations

Differentiate between Repair (corrects the errors in the document) and Restore (recreates the document based on the remaining legible information)

Formats

Input Formats

- PDF 1.x (e.g. PDF 1.4, PDF 1.5, etc.), PDF/A

Target Formats

- PDF 1.x (e.g. PDF 1.4, PDF 1.5, etc.)

Compliance

Standards: ISO 32000 (PDF 1.7)

1.3 Operating Systems

- Windows XP, Vista, 7, 8, 8.1 - 32 and 64 bit

© PDF Tools AG – Premium PDF Technology

- Windows Server 2003, 2008, 2008 R2, 2012, 2012 R2 - 32 and 64 bit

2 Installation

2.1 Overview

The PDF Analysis & Repair Service is configured by the file *pdfrepairsvr.ini*, which needs to be located in the same directory as the executable *pdfrepairsvr.exe*. Before starting the service, the configuration file needs to be adjusted. How this is done is described in the chapter *Editing the Configuration File "pdfrepairsvr.ini"*.

Once configured, the service can be created, started, paused, continued, stopped and deleted via the command line. To use the create and delete functions, administrator permissions are required. To start and stop the service, operator permissions are required.

When the service is running, it processes PDF documents that are copied or moved into watched folders.

The drag and dropped PDF documents are then renamed and moved to the folder *Jobs/*. The renaming gives the PDF a 16 character long timestamp to create unique job tickets. This ensures there are no conflicts with documents that have the same name.

2.2 Installing the Service

1. Download the ZIP archive of the product from your download account at www.pdf-tools.com.
2. Open the ZIP archive.
3. Check the appropriate option to preserve file paths (folder names) and unzip the archive to a local folder (e.g. *C:\program files\pdf-tools*).
4. The unzip process now creates the following subdirectories:
   - Bin: Contains the runtime executable binary code
   - Doc: Contains documentation files

2.3 Editing the Configuration File "pdfrepairsvr.ini"

Before starting the PDF Analysis & Repair Service for the first time, the file *pdfrepairsvr.ini* needs to be modified. Editing this file while the PDF Analysis & Repair Service is running has no impact. The service first needs to be deleted and recreated after the modification. When opening this file with a text editor, it looks like this:

```
[PdfRepairSvr]

AutoDelete=True

Threads=2

Thread1=-w C:\PdfRepairSvr\Analyze -a
```

© PDF Tools AG – Premium PDF Technology

Thread2=-w C:\PdfRepairSvr\Repair

The meaning is as following:

AutoDelete=True This option automatically deletes a PDF file after it is successfully processed. When set to False, the successfully processed file will be copied to the *Succeeded/* subfolder. Failed documents are always copied to the *Failed/* folder.

Threads=2 stands for the total amount of concurrent repair threads. Each thread can have its own assigned settings.

Thread1= sets the options such as output folder, recovery options and other settings for Thread1.

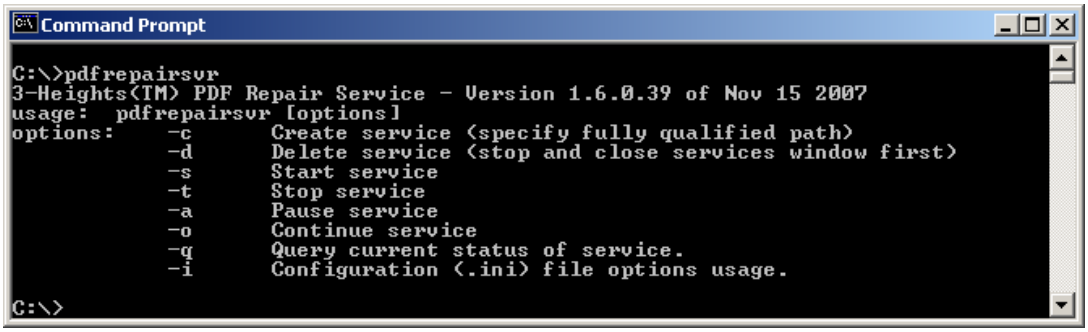
-w C:\PdfRepairSvr\Analyze sets the watched folder for Thread1. The path must be an absolute path. Relative paths or driver letters mapped via the subst command should not be used as they are possibly not recognized by the user under which the PDF Repair Service runs (such as LocalSystem). Network shares are supported when the required permissions are granted.

-a (Option) analyze only option.

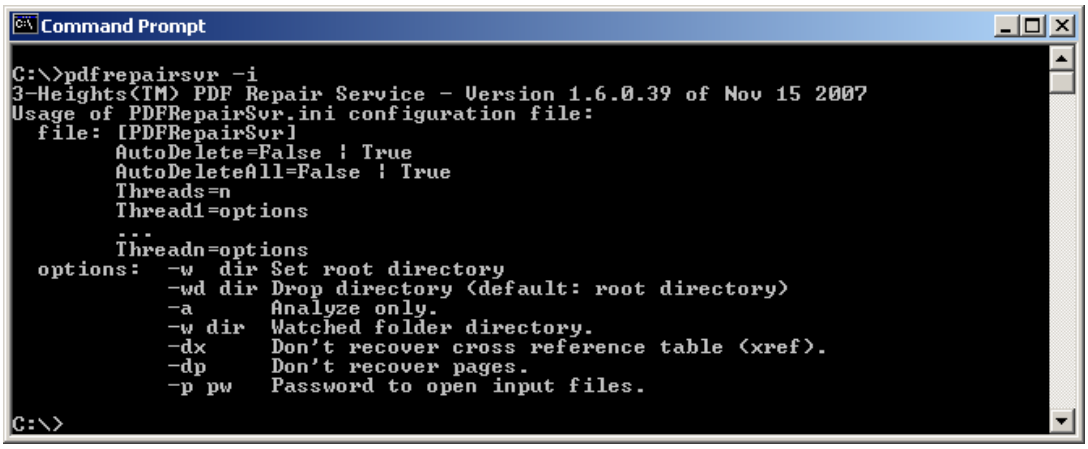
This means if a PDF document is moved to the folder *C:\PdfRepairSvr\Repair* or *C:\PdfRepairSvr\Analyze*, it will be processed by the service. One thread only analyzes, the other also repairs or recovers documents.

Retrieve Information about Available Options

Open a shell and type *pdfrepair* without parameters. This will list the usage, displaying all available options to control the service, as shown in the image below.


Command Prompt
C:\>pdfrepairsvr
3-Heights(TM) PDF Repair Service - Version 1.6.0.39 of Nov 15 2007
usage: pdfrepairsvr [options]
options: -c Create service (specify fully qualified path)
-d Delete service (stop and close services window first)
-s Start service
-t Stop service
-a Pause service
-o Continue service
-q Query current status of service.
-i Configuration (.ini) file options usage.
C:\>

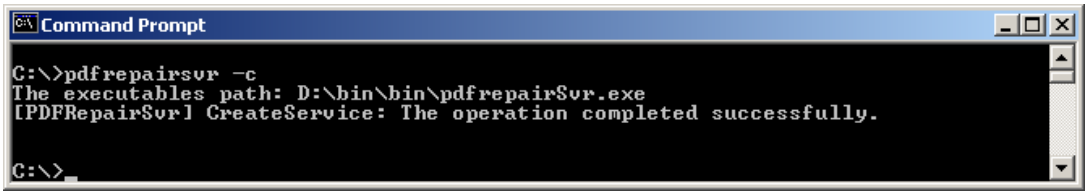
Use the option –i to get a list of all available options for the configuration file:



2.4 Creating and Starting the Service

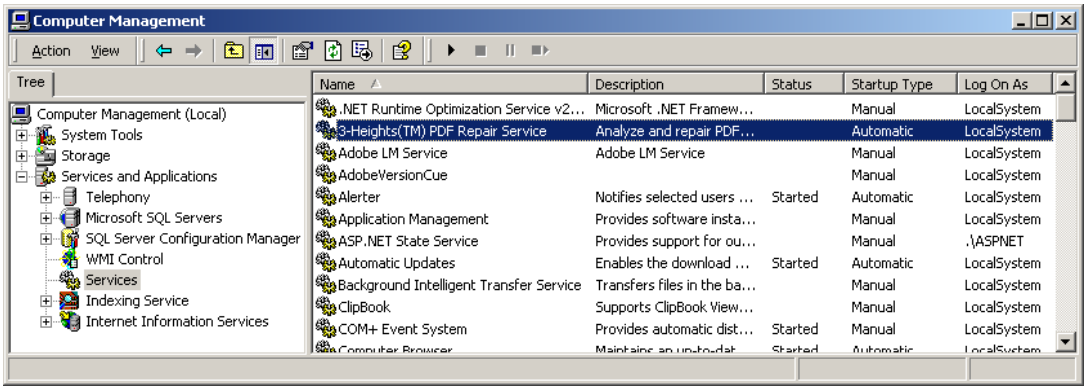
Once the configuration file is completed, the PDF Analysis & Repair Service can be controlled via the command line. To create or delete the service, administrator permissions are required.

1. To create the service, use the option –c. This function will automatically search for the executable path. As shown in the image below:



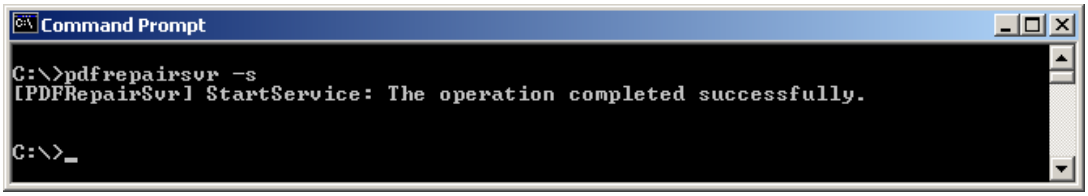
After executing this command, the service is created. It is now visible in the "Computer Management" window under "Services". To open the "Computer Management" window, go to Start -> Control Panel -> Administrative Tools -> Computer Management, or simply right-click the icon "My Computer" on the desktop and select "manage".

The PDF Analysis & Repair Service should then appear as in the image below:

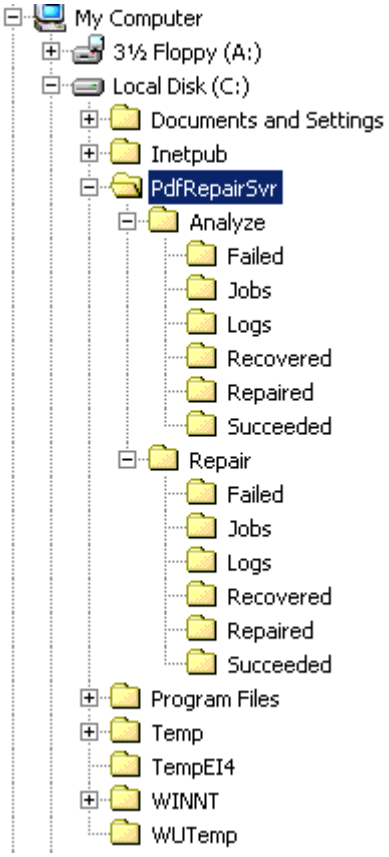

Computer Management
Action View
Tree
Computer Management (Local)
System Tools
Storage
Services and Applications
Telephony
Microsoft SQL Servers
SQL Server Configuration Manager
WMI Control
Services
Indexing Service
Internet Information Services
Name Description Status Startup Type Log On As
.NET Runtime Optimization Service v2... Microsoft .NET Framew... Manual LocalSystem
3-Heights(TM) PDF Repair Service Analyze and repair PDF... Automatic LocalSystem
Adobe LM Service Adobe LM Service Manual LocalSystem
AdobeVersionCue Manual LocalSystem
Alerter Notifies selected users ... Started Automatic LocalSystem
Application Management Provides software insta... Manual LocalSystem
ASP.NET State Service Provides support for ou... Manual .\ASPNET
Automatic Updates Enables the download ... Started Automatic LocalSystem
Background Intelligent Transfer Service Transfers files in the ba... Manual LocalSystem
ClipBook Supports ClipBook View... Manual LocalSystem
COM+ Event System Provides automatic dist... Started Manual LocalSystem

By default the user is set to LocalSystem. After the service is created, the user can be changed. This will be required in a situation where a network share is used as a watched folder and the process needs to run under a user with the appropriate permission rights.

2. After it is created and the user is defined, the service can be started with the option –s from the command line, or directly from the "Service" window. The path can be omitted if the *pdfrepairsvr.exe* is included in the environment variable PATH.



At this time the service creates the watched folders according to the configuration file *pdfrepairsvr.ini* for each thread with six subfolders: *Failed/*, *Job/*, *Logs/*, *Recovered/*, *Repaired/* and *Succeeded/*.



The service is now running and can be used. Here is an example of how PDF documents are processed:

- A PDF document *TestDocument.pdf* is copied into the watched folder *C:\PdfRepairSvr\Repair*.
- It is renamed by adding a 16 Byte time stamp to something like "Job-01C5D099-EDEBD3DA_TestDocument.pdf" and moved to the sub folder *C:\PdfRepairSvr\Repair\Jobs*.
- The next step is to analyze and repair or recover the document:

  If the file is not corrupt, it is moved into the sub folder *Repaired/*.

If the reparation process was successful, a repaired or recovered file is created and available in the subfolder *Repaired/* or *Recovered/* with the same name as the input file in the *Jobs/* folder.

If the reparation process failed, the document is moved from *Jobs/* to *Failed/*.

3. To stop the service, use the option –t. To restart use –s again.
4. To delete the service, first stop it, then use the option –d to delete it.

© PDF Tools AG – Premium PDF Technology

2.5 Error Log Files

There are two types of error log files: By service, and by document.

Log File per Service

This log file contains entries about the service itself. It is created in the same directory as where the executable of the service resides. The log file is locked by the service as long as the service is running. It is named "service-thread-n-log-m", where "n" is the number of the thread and "m" the number of the log file. The number of the log is increased whenever the service is re-started.

- The log file contains general messages such as:

  -Service thread for directory c:\PdfRepairSvr started.

  -Thread for directory c:\PdfRepairSvr terminated.

- Error messages such as:

  * Error 0 while opening file c:\PdfRepairSvr\Repair\Jobs\Job-01C5EE97-6F2667EC_document.pdf.

Log File per Document

The for each document, there is an error log file created. The log file is written to the folder *Log/* and has the same file name as the input document, with extension ".txt". There is no error log available if a document cannot be opened (e.g. is not a PDF document).

There are three types of messages: Information, Warnings, and Errors. They are labelled with the corresponding letters "I", "W" and "E". Here is an example of how a log file could look like:

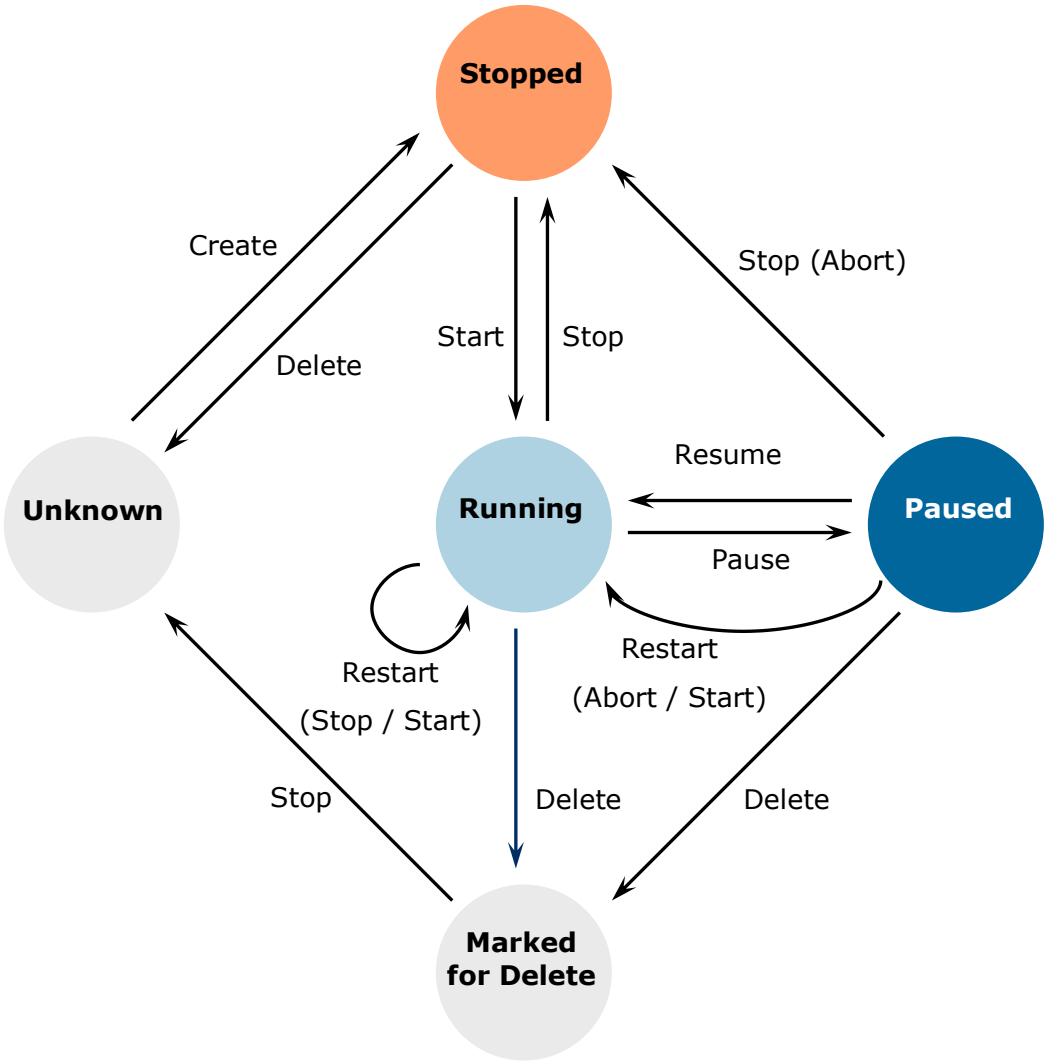
```
Open file.

Analyze Objects.

0x00418018 - W - The generation 0 of reference doesn't match with the
generation 1 of the object.

0x80410028 - E - The "endobj" keyword is missing.

Analyze Pages.

0x80410016 - E - The file is corrupt and cannot be repaired. The file can
possibly be recovered.

Save output file.

Close file.
```

2.6 State Diagram of the Service

The 3-Heights PDF Analysis & Repair Service behaves as described in the state diagram below:

If "Stop" if called when the service is in the state "Paused", the current job is aborted. This means the current page is finished processing, then the job is terminated.

If "Stop" is called when the service is the state "Running", the current job (all pages) is finished. Then the service is stopped.



2.7 Difference between Repairable and Recoverable

Repairable: The input PDF file contains errors which are repairable. These are for example PDF syntax errors. In the repaired PDF output file, these errors are fixed.

Recoverable: The input PDF file is missing data. For example data of an embedded image, which got lost/overwritten when copying the file from one location to another failed partially. When relevant data is lost, the file is irreparable, however the file can be recovered, such that the output PDF is valid according to the PDF Specification.

Example: Assuming an original PDF (file O) is valid.

Somehow the file gets corrupted (-> file C) and file O gets lost.

If file O can be rebuild based on file C, then file C is repairable.

If a new file N can be built based on file C, and file N is a valid PDF and contains (part of) the content of file O, then file C is recoverable.

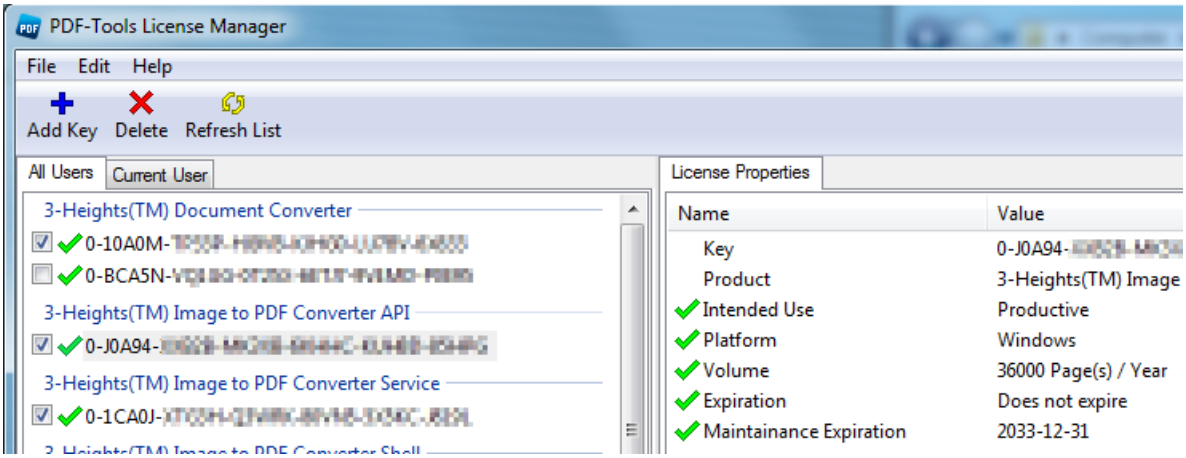
© PDF Tools AG – Premium PDF Technology

3 License Management

For services, there is only one possibility to pass the license key to the application: The license key is installed using the GUI tool (Graphical user interface). This is the easiest way if the licenses are managed manually.

3.1 Graphical License Manager Tool

The GUI tool *LicenseManager.exe* is located in the *bin* directory of the product kit.



List all installed license keys

The license manager always shows a list of all installed license keys in the left pane of the window. This includes licenses of other PDF Tools products.

The user can choose among:

- Licenses available for all users. Administrator rights are needed for modifications.
- Licenses available for the current user only.

Add and delete license keys

License keys can be added or deleted with the "Add Key" and "Delete" buttons in the toolbar.

- The "Add key" button installs the license key into the currently selected list.
    - Note: Services run by default under the LOCAL SERVICE user, not under the current user.
- The "Delete" button deletes the currently selected license keys.

Display the properties of a license

If a license is selected in the license list, its properties are displayed in the right pane of the window.

Select between different license keys for a single product

More than one license key can be installed for a specific product. The checkbox on the left side in the license list marks the currently active license key.

3.2 Platform considerations

Windows

The license keys are stored in the registry:

- HKLM\Software\PDF Tools AG (for all users)
- HKCU\Software\PDF Tools AG (for the current user)

4 Reference Manual

4.1 Service Control Commands

These options are used to control the service. The *create* and *delete* functions require administrator rights. The *start, pause, continue* and *stop* functions require operator rights.

To create and start the service, use the sequence:

- –c
- optionally adjust logon
- -s

To stop and shut down the service use the sequence:

- -t
- -d

When the configuration file is changed, stop and start the service.

-c Create Service

The 3-Heights PDF Analysis & Repair Service is created using the option –c.

```
C:\> pdfrepairsvr -c
```

Important: It is essential that pdfrepairsvr.exe be on a **not-mapped** drive.

-d Delete Service

The PDF Analysis & Repair Service can be deleted with the option –d.

```
C:\> pdfrepairsvr –d
```

© PDF Tools AG – Premium PDF Technology

-s Start Service

Once created, the service can be started with the option –s.

```
C:\> pdfrepairsvr –s
```

-t Stop Service

To stop the service, use the option –t.

```
C:\> pdfrepairsvr –t
```

If "stop" is called while the service is "running", the current job (all pages) will be finished, after that the service is stopped.

If the service was "paused" before calling "stop", the current page while be finished processing. After that page, the job is aborted.

-o Continue Service

This option resumes the service.

```
C:\> pdfrepairsvr –o
```

-a Pause Service

This option pauses the service.

```
C:\> pdfrepairsvr –a
```

-q Query Current Sstatus

This options returns the current status of the service.

```
C:\> pdfrepairsvr -q
The service starts automatically during system startup.
The service is running.
[PDFRepairSvr] QueryService: The operation completed successfully.
```

-x Run as Executable

With this option, the PDF Repair runs as an executable instead of as a Windows service. It provides the same functionality as long as the exe is running.

```
C:\> pdfrepairsvr –x
```

4.2 Configuration Options

Analyzing and repairing or recovering PDF documents is a complex task. To increase to processing speed, certain analyze and recovery options can be disabled.

Configuration File *Pdf2PdfSvr.ini*

The configuration file defines the setting for the watched folders. It is read upon starting the service.

`[PdfRepairSvr]`	required
`Autodelete=...`	optional, `true` or `false`
`AutodeleteAll=...`	optional, `true` or `false`
`LogPath=...`	optional, if used must be a path like `C:\mypath\log` or the keyword `EventLog`
`PollingInterval=...`	optional, value in milliseconds, default `1000`
`JobPrefix=...`	optional, `true` or `false`
`Threads=n`	required
`Thread1=-w ...`	required
`Thread2=-w ...`	
`...`	
`Threadn=...`	There must be exact as many threads as defined in Threads=n.

Example:

```
[PdfRepairSvr]
Autodelete=true
LogPath=EventLog
JobPrefix=false
Threads=2
Thread1=-w C:\PdfRepairSvr\Analyze -a
Thread2=-w C:\PdfRepairSvr\Repair
```

-a Analyze Only

When using this option, the processed input files are analyzed only and a log file is generated. There is no output created.

-b Set the Rebuild Options

This options controls what parts of the PDF are to be repaired. Available options are:

1: Rebuild streams
2: Rebuild fonts
4: Convert CFF fonts to Type1 fonts.

© PDF Tools AG – Premium PDF Technology

If 4 is applied, the compressed fonts are decompressed, this can potentially lead to an increase of the file size.

If multiple options are to be selected, add the values.

To repair the file and rebuild all use:

```
-b 7
```

-dp Do Not Recover Pages

If pages are not part of the page tree (loose pages), they will be recovered and added at the end of the document. If the option "-dp" is selected, loose pages not be recovered and will be left out of the recovered document.

-dx Do Not Recover Cross Reference Table

With this option a corrupt XREF table will not be repaired or recovered. This option is useful if it is taking too long to process a document, since repairing the cross-reference table is very time consuming.

-p Read an Encrypted PDF File

A PDF document that has a user password (the password to open the document) can only be read when either the user or the owner password is provided. The password can be provided using the option -p followed by the password.

```
-p mypassword
```

-w Set the Watched Folder

Use the option –w to define the path of the watched folder. This path should not contain mapped drives, since other users (such as LocalSystem) do not recognize them. This parameter must always be the first parameter of a thread.

```
-w C:\ PdfRepairSvr\Analyze
```

Note that the service supports path lengths including file name of up to 258 characters. This includes the 21 characters of the job ticket. If a file name exceeds this value, its file name is truncated at the end of the file name and before the file extension. It is therefore suggested that watched folder names are kept reasonably short.

-wd Set the Drop-In Folder

By default the drop-in folder is equal to the folder defined as watched folder using the option –w. If the input files should be taken from a different folder, this can be configured using –wd. All folders created by service including the output folder are at the directory defined by –w.

```
-wd C:\SomePath\DropIn
```

© PDF Tools AG – Premium PDF Technology

-wfs Process only files with certain extensions

By default, the service tries to process all files dropped into the drop-in folder, regardless of the extension. With this option, the processing can be restricted to a set of known file extension.

Example: *Restrict the processing to JPEG and PNG files.*

```
-wfs .jpeg.jpg.png
```

-wfi Ignore files with certain extensions

By default, the service tries to process all files dropped into the drop-in folder, regardless of the extension. With this option, files with certain file extensions ca be ignored.

Example: *Ignore temporary files.*

```
-wfi .temp.tmp
```

Autodelete Successfully Processed Files

When a repair job succeeded, the PDF document will be moved from the folder *Jobs/* to the folder *Repaired/* or *Recovered/*. To automatically delete the input file after it has be successfully repaired or recovered, the value "AutoDelete" can be set to True in the control file *pdfrepairsvr.ini*. The documents which cannot be repaired nor recovered are still copied to the subfolder *Failed/*.

```
[PdfRepairSvr]
AutoDelete=True
```

To delete failed documents as well use the following setting:

```
[PdfRepairSvr]
AutoDeleteAll=True
```

Job Number Prefix

At the time when a document is copied from the watched folder to the *Jobs/* sub folder, it is renamed and added a 21 character prefix containing a time stamp: Job-{8 digits}-{8 digits}_ Something like: "Job-01C61DD4-E72E1BCE_". The job number prefix ensures the same document (name) can be processed multiple times.

Adding the prefix can be prevented with the following line in the configuration file:

```
[PdfRepairSvr]
JobPrefix=False
```

Logpath

Log-messages created by the service are by default written to the sub-directory *log*. To alter the directory, add a line similar as shown below to the configuration file:

```
[Pdf2PdfSvr]
```

```
LogPath=C:\path\log
```

Messages created by the service can be added to the system's application event log instead of written to a log file. This is achieved by adding the following line to the configuration file:

```
[Pdf2PdfSvr]
LogPath=EventLog
```

The system's application log event will then log messages similar as shown below:

- CreateService: The operation completed successfully.
- StartService: The operation completed successfully.

Note that the messages are only fully accessible while the service is created. Otherwise a message as shown below is displayed:

- The description for Event ID (1) in Source (Pdf2PdfSvr) cannot be found. The local computer may not have the necessary registry information or message DLL files to display messages from a remote computer. The following information is part of the event: DeleteService: The operation completed successfully.

Polling Interval

The polling interval defines the time in milliseconds that the polling-thread pauses between two polls. The time passing until the same watched folder is polled again (maximum pick-up time) is: The value of PollingInterval plus the actual time it takes to poll all watched folders. The higher the polling interval, the lower the network traffic, and the longer it takes until documents are picked up.

Suggested values for the polling intervals are 1000 to 10000 milliseconds.

```
[PdfRepairSvr]
PollingInterval=5000
```

5 Troubleshooting

5.1 The Repair Time Takes Too Long

Try excluding the cross-reference table (switch **–dx**). This will speed up the repair time.

5.2 The File Cannot be Repaired

If the switch **–dx** has been selected and the cross-reference table is corrupt, the file cannot be repaired. Ensure the switch **–dx** is not set.

© PDF Tools AG – Premium PDF Technology